WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

A PROBABILISTIC ANALYSIS OF THE MULTIKNAPSACK VALUE FUNCTION

by

M. Meanti (*)

A.H.G. Rinnooy Kan (**)

L. Stougie (***)

C. Vercellis (*)

Working paper # 1611-84

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

A PROBABILISTIC ANALYSIS OF THE MULTIKNAPSACK VALUE FUNCTION

by

M. Meanti (*)

A.H.G. Rinnooy Kan (**)

L. Stougie (***)

C. Vercellis (*)

Working paper # 1611-84

September 1984

(*) Dipartimento di Matematica, University di Milano, Italy

(**) Econometric Institute, Erasmus University Rotterdam, The Netherlands/Sloan School M.I.T., Cambridge, Massachusetts, U.S.A.

(***) Center for Mathematics and Computer Science, Amsterdam, The Netherlands

---

This research was partially supported by NSF Grant ECS-831-6224, and MPI Project "Matematica computazionale".

1. Introduction

The *Multi Knapsack* (MK) problem is defined as follows. Let each item j (j=1,2,...n) require a certain amount of space a_{ij} of knapsack i (i=1,2,...m) and let c_j denote the profit of having item j in the knapsacks. Let x_j be a zero-one variable, taking value one if item j is included in the knapsacks, and zero otherwise. The capacities of the knapsacks are denoted by $b_1, b_2, \ldots b_m$.

This leads to the following optimization problem:

$$\max \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \sum_{j=1}^{n} a_{ij} x_j \leq b_i \qquad (i=1,2,\ldots m) \qquad \text{(MK)}$$

$$x_j \in \{0,1\} \qquad (j=1,2,\ldots n).$$

MK has been used to model problems in the areas of scheduling and capital budgeting /10/. The problem is known to be NP-hard /6/; it is a generalization of the standard knapsack problem (m=1). MK can be solved by a polynomial approximation scheme /5/, but a fully polynomial one cannot exist unless P=NP /9/.

In this paper, we are interested in analysing the behaviour of the optimal value of MK with respect to changing capacities of the knapsacks. More specifically, we would like to obtain an expression that represents the optimal value as a function of b_i (i=1,2,...m) over a range of problem instances.

A stochastic model for the problem parameters c_j, a_{ij} (i=1,2,...m, j=1,2,...n) will show that the sequence of optimal values, properly normalized, converges with probability one (wp1) to a function of the b_i's as the number n of items tends to infinity while the number m of knapsacks is fixed.

A number of probabilistic analyses of approximate algorithms for the knapsack problem (m=1) have been conducted in the past (/1/, /2/, /7/, /11/). So far, the optimum value has not been asymptotically characterized as a function of the capacity b of the knapsack. A similar comment applies to the probabilistic analysis developed in /5/ for MK under a stochastic model less general than the one we consider here. In fact, our only requirement will be absolute continuity of the distributions involved; the results are unusually robust.

The main result in Section 3 is the asymptotic characterization of the optimum value as a function of b_i (i=1,2,...m). This characterization is obtained by showing that the optimal values of the LP-relaxation of MK (denoted by MKLP) have a regular limiting behaviour holding wp1. The fact that the optimal values of MK and MKLP are asymptotic to each other follows from a result in Section 2.

Three other relevant results are obtained. To characterize the integer optimal solution itself, we shall prove that the zero-one vectors in a certain sequence are optimal solutions of MK infinitely often (i.o.). We also show that, relative to the optimal integer solution, at least one constraint has a slack variable whose expected value equals zero i.o. . Finally, we prove that the sequence of optimal dual multipliers of MKLP also has a regular asymptotic behaviour, converging to a limit wp1.

In Section 4, the results obtained are applied to problem MK for the cases that m=1 and m=2 and for specific distributions of the parameters so as to obtain explicit formulae for the optimal value of MK. They are depicted in Figures 1 and 3.

Finally, in Section 5, an approximation algorithm is proposed. It generalizes the greedy heuristic for MK (cf. /3/) and will be shown to be asymptotically optimal wp1.

2. Stochastic model and upper and lower bounds

Let us assume that the profit parameters c_j $(j=1,2,\ldots n)$ are independent identically distributed (i.i.d.) random variables (r.v.), defined over a bounded interval in $\mathbb{R}$ with common distribution F_c. Analogously, for each $i=1,2,\ldots,m$, the requirement coefficients a_{ij} $(j=1,2,\ldots,n)$ are assumed to be i.i.d. r.v.'s with common distribution F_i over a bounded interval.

In the sequel, r.v.'s will be underlined. Without loss of generality the interval on which $\underline{c}_j$, $\underline{a}_{ij}$ are defined will be assumed to be $[0,1]$. We further assume that the distributions F_c, F_i $(i=1,2,\ldots,m)$ are absolutely continuous and that the corresponding densities f_c, f_i $(i=1,2,\ldots,m)$ are continuous and strictly positive over $(0,1)$.

It is reasonable to assume that the capacities b_i grow proportionally to the number of items. Specifically, let $b_i=n\beta_i$ $(i=1,2,\ldots,m)$ for $\beta \in V = \{\beta\colon 0<\beta_i<E[\underline{a}_{i1}],\ (i=1,2,\ldots,m)\}$ As remarked in /11/ , the i-th constraint would tend to be redundant if $\beta_i \geq E[\underline{a}_{i1}]$, in the sense that it would be satisfied even if all items are included, with probability tending to one as $n\to\infty$.

Define

$$\underline{z}_n^I=\max\{ \sum_{j=1}^n \underline{c}_j x_j \mid \sum_{j=1}^n \underline{a}_{ij} x_j \leq n\beta_i \ (i=1,2,\ldots,m), x_j \in \{0,1\} (j=1,2,\ldots n)\}$$

to be the optimal value of MK and

$$\underline{z}_n^{LP}=\max\{ \sum_{j=1}^n \underline{c}_j x_j \mid \sum_{j=1}^n \underline{a}_{ij} x_j \leq n\beta_i \ (i=1,2,\ldots,m),\ 0\leq x_j \leq 1 \ (j=1,2,\ldots n)\}$$

to be the optimal value of MKLP.

By definition

$$z_n^{LP} \geqslant z_n^I \tag{2.1}$$

for each realization of the random parameters.

To bound $\underline{z}_n^I$ from below the following lemma can easily be established :

<u>LEMMA 2.1</u> : $z_n^I \geqslant z_n^{LP} - m.$

<u>Proof</u> : The optimal solution of MKLP has at most m basic variables that are fractional. Rounding down these values yields a feasible integer solution which is a lower bound on z_n^I . The value corresponding to this rounddown solution is given by z_n^{LP} decreased by at most $m \cdot \max\{c_j\} = m.$ ■

3. Asymptotic characterization of the optimal value

In order to derive a function of b_i $(i=1,2,\dots,m)$ to which $\underline{z}_n^{LP}$ is asymptotic wp1, we consider the *Lagrangean relaxation* of the problem MKLP, defined as

$$\underline{w}_n(\lambda) = \max \left\{ \sum_{i=1}^{m} \lambda_i b_i + \sum_{j=1}^{n} \left(\underline{c}_j - \sum_{i=1}^{m} \lambda_i \underline{a}_{ij} \right) x_j \,\middle|\, 0 \leq x_j \leq 1 \ (j=1,2,\dots,n) \right\}$$

It is well known that, for every realization of the stochastic parameters, the function $w_n(\lambda)$ is convex over the region $\lambda \geq 0$. Moreover,

$$\min_{\lambda \geq 0} w_n(\lambda) = z_n^{LP} \quad . \tag{3.1}$$

Let $\underline{\lambda}_n^*$ be a vector of multipliers minimizing $\underline{w}_n(\lambda)$. Define the r.v.'s

$$\underline{x}_j^L(\lambda) = \begin{cases} 1 & \text{if } \underline{c}_j - \sum_{i=1}^{m} \lambda_i \underline{a}_{ij} \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (j=1,2,\dots n).$$

Then

$$\underline{w}_n(\lambda) = \sum_{i=1}^{m} \lambda_i b_i + \sum_{j=1}^{n} \left(\underline{c}_j - \sum_{i=1}^{m} \lambda_i \underline{a}_{ij} \right) \underline{x}_j^L(\lambda).$$

Define also

$$\underline{L}_n(\lambda) = \frac{1}{n} \underline{w}_n(\lambda) = \sum_{i=1}^{m} \lambda_i \beta_i + \frac{1}{n} \sum_{j=1}^{n} \left(\underline{c}_j - \sum_{i=1}^{m} \lambda_i \underline{a}_{ij} \right) \underline{x}_j^L(\lambda).$$

We first prove a preliminary lemma that establishes the almost sure convergence of $\underline{L}_n(\lambda)$ to a non-random function $L(\lambda)$, defined by

$$L(\lambda) = \sum_{i=1}^{m} \lambda_i \beta_i + E\ \underline{cx}^L(\lambda) - \sum_{i=1}^{m} \lambda_i\ E\ \underline{a}_i\underline{x}^L(\lambda) .$$

Here $E\ \underline{cx}^L(\lambda)$ and $E\ \underline{a}_i\underline{x}^L(\lambda)$ are used to denote respectively the common values of $E\left[\underline{c}_j\underline{x}_j^L(\lambda)\right]$ and $E\left[\underline{a}_{ij}\underline{x}_j^L(\lambda)\right]$ $(j=1,2,\ldots n)$. We shall also write $E\ \underline{ax}^L(\lambda)$ for the vector $(E\ \underline{a}_1\underline{x}^L(\lambda),\ E\ \underline{a}_2\underline{x}^L(\lambda)\ \ldots,\ E\ \underline{a}_m\underline{x}^L(\lambda))$.

We observe that the r.v.'s $\{\underline{c}_j\underline{x}_j^L(\lambda)\}$ $(j=1,2,\ldots n)$ are i.i.d., as well as the r.v.'s $\{\underline{a}_{ij}\underline{x}_j^L(\lambda)\}$ $(j=1,2,\ldots n)$ for any $i \in \{1,2,\ldots m\}$. This property will be used throughout the paper, sometimes without being explicitly mentioned. In particular, it can be used to apply the strong law of large numbers to the sequence $\underline{L}_n(\lambda)$ and obtain the following result.

<u>LEMMA 3.1</u>: For every $\lambda \geq 0$, $\underline{L}_n(\lambda) \xrightarrow{wp1} L(\lambda)$.■

The next two lemmas, whose proofs are given in the Appendix, describe some properties of the function $L(\lambda)$ which will turn to be useful in the sequel.

<u>LEMMA 3.2</u>: $L(\lambda)$ is twice continuously differentiable and strictly convex. For each $\beta \in V$ it has a unique non-zero minimum $\lambda^* = \lambda^*(\beta)$ over the region $\lambda \geq 0$. The gradient is given by

$$\nabla L(\lambda) = \beta - E\ \underline{ax}^L(\lambda) .■$$

<u>LEMMA 3.3</u>: For λ^* the following conditions are satisfied for $i=1,2,\ldots m$:

(i) $\lambda_i^*(\beta_i - E\,\underline{a}_i \underline{x}^L(\lambda^*)) = 0$

(ii) $E\,\underline{a}_i \underline{x}^L(\lambda^*) \leq \beta_i$. ■

If we can prove that $\underline{L}_n(\underline{\lambda}_n^*) \xrightarrow{wp1} L(\lambda^*)$, then through (3.1) $L(\lambda^*)$ would provide an asymptotic characterization of $\frac{1}{n}\,\underline{z}_n^{LP}$.

To establish this convergence, we have to strengthen the result in Lemma 3.1 from pointwise to uniform convergence wp1. To do this, we shall apply a theorem from /12/ which states that pointwise convergence of a sequence of convex functions on a compact subset of their domain implies uniform convergence on this subset to a function that is also convex.

First, we have to show that the minima $\{\underline{\lambda}_n^*\}$ (n=1,2,...) and λ^* are contained in a compact subset $S \subseteq R_+^m$. In fact, we shall show that the minimization of the functions $\underline{L}_n(\lambda)$ (n=1,2,...) and $L(\lambda)$ over R_+^m is equivalent to their minimization over the set

$$S = \{\lambda \mid \sum_{i=1}^{m} \beta_i \lambda_i \leq 1,\ \lambda_i \geq 0 \quad (i=1,2,\ldots,m)\} .$$

LEMMA 3.4: For every realization of the stochastic parameters, the functions $L(\lambda)$ and $L_n(\lambda)$ (n=1,2,...) attain their minimum within the set S.

Proof: it is easy to see that

$$L_n(0) = \frac{1}{n} \sum_{j=1}^{n} c_j \leq 1$$

For any other value of λ it holds that

$$L_n(\lambda) = \lambda\ \beta + \frac{1}{n} \sum_{j=1}^{n} (c_j - \sum_{i=1}^{m} \lambda_i a_{ij})\ x_j^L(\lambda) \geq \lambda\ \beta\ .$$

Hence, for every λ with $\lambda\beta > 1$, we have

$$L_n(\lambda) > 1 \geq L_n(0) \quad .$$

Convexity of $\underline{L}_n(\lambda)$ implies, together with the inequalities above, that $\lambda_n^* \in S$, $n=1,2,\ldots$ .The same arguments applied to $L(\lambda)$ show that $\lambda^* \in S$. ∎

As a direct application of Rockafellar's theorem 10.8 /12/, we then obtain the following result.

<u>LEMMA 3.5</u> : $\underline{L}_n(\lambda) \xrightarrow{wp1} L(\lambda)$ uniformly on S.

We are now in a position to prove the required result.

<u>THEOREM 3.1</u>: $\underline{L}_n(\underline{\lambda}_n^*) \xrightarrow{wp1} L(\lambda^*)$.

<u>Proof</u>: Uniform convergence wp1 on S can be written as

$$\Pr\{\forall\varepsilon>0 \quad \exists n_o : \forall n \geq n_o \quad \sup_{\lambda \in S} |\underline{L}_n(\lambda) - L(\lambda)| < \varepsilon\} = 1 \quad . \tag{3.2}$$

It is easy to see that

$$|\underline{L}_n(\underline{\lambda}_n^*) - L(\lambda^*)| \leq \sup_{\lambda \in S} |\underline{L}_n(\lambda) - L(\lambda)| \quad . \tag{3.3}$$

The combination of (3.2) and (3.3) yields

$$\Pr\{\forall\varepsilon>0 \quad \exists n_o : \forall n \geq n_o \quad |\underline{L}_n(\underline{\lambda}_n^*) - L(\lambda^*)| < \varepsilon\} = 1$$

which proves the required result. ∎

From (3.1) and Theorem 3.1 it follows that

$$\frac{1}{n}\,\underline{z}_n^{LP} \xrightarrow{wp1} L(\lambda^*) \,. \tag{3.4}$$

Moreover, we know from Lemma 2.1 that, for each realization,

$$\frac{1}{n}\,z_n^{LP} - \frac{m}{n} \leq \frac{1}{n}\,z_n^{I} \leq \frac{1}{n}\,z_n^{LP} \,. \tag{3.5}$$

Combining (3.4) and (3.5) one can easily establish the following theorem.

<u>THEOREM 3.2</u>: $\frac{1}{n}\,\underline{z}_n^{I} \xrightarrow{wp1} L(\lambda^*)$. ■

This latter result gives the required asymptotic characterization of the optimum of MK, holding almost surely. Observe that $L(\lambda^*)$ is actually a function of the right-hand side b_i $(i=1,2,\ldots m)$ implicitly defined by minimization of $L(\lambda)$ over S. The problem of deriving a closed-form expression for this function, or at least of evaluating it numerically for different values of its arguments b_i $(i=1,2,\ldots m)$ will be considered in the subsequent Section 4 for specific distributions $F_c, F_i (i=1,\ldots,m)$

Whereas Theorem 3.2 is concerned with the behaviour of the optimum value $\underline{z}_n^I$ of MK, the following result says something about the corresponding optimal solution.

<u>THEOREM 3.3</u> : The vector $\underline{x}^L(\lambda^*)=(\underline{x}_1^L(\lambda^*),\ \underline{x}_2^L(\lambda^*)\ldots\underline{x}_n^L(\lambda^*))$ is optimal infinitely often (i.o.).

<u>Proof</u> : Define the following events: $Q_n=\{\underline{x}^L(\lambda^*)$ is not optimal$\}$, $R_n=\{\underline{x}^L(\lambda^*)$ is suboptimal$\}$ and $T_n=\{\underline{x}^L(\lambda^*)$ is not feasible$\}$.

By definition, we have to show that the series $\Sigma \Pr\{Q_n\}$ is convergent. Obviously,

$$\Pr\{Q_n\} \leq \Pr\{R_n\} + \Pr\{T_n\} \quad .$$

We have

$$\Pr\{R_n\} \leq \Pr\{\ \exists \varepsilon :\ \frac{1}{n}\sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\lambda^*) < \frac{1}{n}\ \underline{z}_n^{LP} - \varepsilon\} =$$

$$= \Pr\{\frac{1}{n}\ |\ \underline{z}_n^{LP} - \sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\lambda^*)| > \varepsilon\} \quad .$$

Since $\frac{1}{n}\ (\sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\lambda^*) - \underline{z}_n^{LP}) \xrightarrow{wp1} 0$, the series $\Sigma \Pr\{R_n\}$ converges for every $\varepsilon>0$. Moreover,

$$\Pr\{T_n\} = \Pr\{\ \exists i, \varepsilon>0:\ \beta_i - \frac{1}{n}\sum_{j=1}^{n} \underline{a}_{ij} \underline{x}_j^L(\lambda^*) < -\varepsilon\}$$

$$\leq m \Pr\{\beta_i - \frac{1}{n}\sum_{j=1}^{n} \underline{a}_{ij} \underline{x}_j^L(\lambda^*) < -\varepsilon\},$$

so that the series $\Sigma \Pr\{T_n\}$ is also convergent, since from Lemma 3.3

$$\beta_i - \frac{1}{n}\sum_{j=1}^{n} \underline{a}_{ij} \underline{x}_j^L(\lambda^*) \xrightarrow{wp1} \beta_i - E\ \underline{a}_i \underline{x}^L(\lambda^*) \geq 0 \ (i=1,2,\ldots m). \blacksquare$$

As a consequence of Lemmas 3.2, 3.3 and Theorem 3.4 at least one constraint will be binding in expectation at the integer optimum i.o., in the sense that the expected value of its slack equals zero i.o. .

However, this property is not true in general for all constraints (cf. our comments for the case m=2 in Section 4).

To conclude this section a stronger version of Theorem 3.1 will be proved, which extends the convergence result from the sequence of minima of the Lagrangean $\underline{L}_n(\lambda)$ to the sequence of their arguments, showing that $\underline{\lambda}_n^* \xrightarrow{wp1} \lambda^*$. This property indicates regular asymptotic behaviour of the sequence of optimal dual multipliers $\underline{\lambda}_n^*$. Moreover, it will be useful in developing the approximation algorithm proposed in Section 5.

<u>THEOREM 3.4</u> : $\underline{\lambda}_n^* \xrightarrow{wp1} \lambda^*$.

<u>Proof</u>: A Taylor series of $L(\lambda)$ around λ^* can be written as

$$L(\underline{\lambda}_n^*)-L(\lambda^*) = (\underline{\lambda}_n^*-\lambda^*)\ \nabla L(\lambda^*) + \frac{1}{2}(\underline{\lambda}_n^*-\lambda^*)H(\lambda_n)(\underline{\lambda}_n^*-\lambda^*) \qquad (3.6)$$

where $H(\lambda_n)$ is the Hessian matrix of $L(\lambda)$ evaluated at a point λ_n lying on the line segment $[\underline{\lambda}_n^*, \lambda^*]$ and therefore belonging to the convex set S.

Positive definiteness of $H(\lambda)$ for every λ (see the Appendix) implies strict positivity of its eigenvalues. Let $\alpha(\lambda)$ denote the smallest eigenvalue of $H(\lambda)$, and let $\alpha=\inf_{\lambda\in S} \alpha(\lambda)$; note that the continuity of $\alpha(\lambda)$ implies that $\alpha>0$.

It can be easily shown that, for any n,

$$\frac{1}{2}(\underline{\lambda}_n^*-\lambda^*)\ H(\lambda_n)(\underline{\lambda}_n^*-\lambda^*) \geq \frac{\alpha}{2}\ \|\underline{\lambda}_n^*-\lambda^*\|^2 \quad . \qquad (3.7)$$

Moreover, due to Lemma 3.3, we have

$$-\lambda^{*}\nabla L(\lambda^{*}) = 0 \quad , \qquad \underline{\lambda}_n^{*}\nabla L(\lambda^{*}) \geq 0 \ . \tag{3.8}$$

Combining (3.6), (3.7) and (3.8), we obtain that

$$L(\underline{\lambda}_n^{*})-L(\lambda^{*}) \geq \frac{\alpha}{2} \ || \underline{\lambda}_n^{*} - \lambda^{*}||^{2}.$$

Consequently, to prove convergence of $\underline{\lambda}_n^{*}$ to λ^{*}, it remains to show that $L(\underline{\lambda}_n^{*}) \xrightarrow{wp1} L(\lambda^{*})$.
Indeed,

$$|L(\underline{\lambda}_n^{*}) - L(\lambda^{*})| \leq |L(\underline{\lambda}_n^{*})-\underline{L}_n(\underline{\lambda}_n^{*})|+|\underline{L}_n(\underline{\lambda}_n^{*})-L(\lambda^{*})| \ ,$$

and $|L(\underline{\lambda}_n^{*})-\underline{L}_n(\underline{\lambda}_n^{*})| \xrightarrow{wp1} 0$ by Lemma 3.5, whereas

$|\underline{L}_n(\underline{\lambda}_n^{*})-L(\lambda^{*})| \xrightarrow{wp1} 0$ by Theorem 3.1.■

4. Particular cases: m=1, m=2

The actual computation of E $\underline{c}\underline{x}^L(\lambda)$ and E $\underline{a}_i\underline{x}^L(\lambda)$ requires a choice for a specific distribution for the stochastic parameters of MK. In this section, we assume that the profit coefficients $\underline{c}_j$ as well as the requirements $\underline{a}_{ij}$ are uniformly distributed over $[0,1]$. This assumption is essentially the same as in / 5/, / 7/, /11/.

We first present and discuss the results concerning MK with one constraint. In this case, straightforward calculations lead to the following formulae:

$$E\ \underline{a}\underline{x}^L(\lambda) = \begin{cases} \dfrac{1}{2} - \dfrac{\lambda}{3} & \text{if } \lambda \leq 1 \\[2ex] \dfrac{1}{6\lambda^2} & \text{if } \lambda > 1 \end{cases}$$

$$E\ \underline{c}\underline{x}^L(\lambda) = \begin{cases} \dfrac{1}{2} - \dfrac{\lambda^2}{6} & \text{if } \lambda \leq 1 \\[2ex] \dfrac{1}{3\lambda} & \text{if } \lambda > 1 \end{cases}$$

$$L(\lambda) = \begin{cases} \lambda\left(\beta - \dfrac{1}{2} + \dfrac{\lambda}{3}\right) + \dfrac{1}{2} - \dfrac{\lambda^2}{6} & \text{if } \lambda \leq 1 \\[2ex] \lambda\left(\beta - \dfrac{1}{6\lambda^2}\right) + \dfrac{1}{3\lambda} & \text{if } \lambda > 1 \end{cases}$$

$$\lambda^* = \begin{cases} \dfrac{1}{\sqrt{6\beta}} & \text{if } 0 < \beta < \dfrac{1}{6} \\[2ex] \dfrac{3}{2} - 3\beta & \text{if } \dfrac{1}{6} \leq \beta < \dfrac{1}{2} \end{cases}$$

Therefore, by Theorem 3.2,

$$\frac{1}{n}\, \underline{z}_n^I \xrightarrow{wp1} L(\lambda^*) = \begin{cases} \sqrt{\frac{2}{3}\,\beta} & \text{if } 0<\beta<\frac{1}{6} \\ -\frac{3}{2}\beta^2 + \frac{3}{2}\beta + \frac{1}{8} & \text{if } \frac{1}{6}\leq\beta<\frac{1}{2} \end{cases} .$$

The graph of $L(\lambda^*)$ as a function of β is shown in Figure 1.

Notice that in accordance with Lemmas 3.2 and 3.3, $E\,\underline{ax}^L(\lambda^*)-\beta=0$.

In the case that m=2, $E\,\underline{cx}^L(\lambda)$, $E\,\underline{a}_1\underline{x}^L(\lambda)$ and $E\,\underline{a}_2\underline{x}^L(\lambda)$ take different forms over different regions of the $\mathbb{R}_+^2$ plane. We will describe directly the corresponding functions $\lambda^*=\lambda^*(\beta)$ defined on the (β_1, β_2) plane, referring to Figure 2.

Define the following regions:

$$A = \{(\beta_1,\beta_2) : \beta_1\geq\beta_2,\ \ \beta_2\geq 24\beta_1^2\}$$

$$B = \{(\beta_1,\beta_2): \beta_2\leq\frac{8}{3}\beta_1^2,\ \beta_2\leq\frac{1}{6},\ \beta_1<\frac{1}{2}\}$$

$$C = \{(\beta_1,\beta_2): \beta_2>\frac{1}{6},\ \beta_2\leq\frac{4}{3}\beta_1 - \frac{1}{6},\ \beta_1<\frac{1}{2}\}$$

$$D = \{(\beta_1,\beta_2): \beta_1\geq\beta_2,\ \beta_2>\frac{4}{3}\beta_1 - \frac{1}{6},\ \beta_1+\beta_2\geq\frac{5}{12}\}$$

$$E = \{(\beta_1,\beta_2): \beta_1\geq\beta_2,\ \beta_2>\frac{8}{3}\beta_1^2,\ \beta_1+\beta_2<\frac{5}{12}\} .$$

The values of λ^* and $L(\lambda^*)$ in the corresponding five regions where $\beta_1<\beta_2$ can be obtained by exchanging β_1 with β_2 and viceversa in the formulae given below.

Region A:

$$\lambda_1^* = \sqrt[3]{\frac{\beta_2}{24\beta_1^2}} \quad , \quad \lambda_2^* = \sqrt[3]{\frac{\beta_1}{24\beta_2^2}}$$

$$L(\lambda^*) = \frac{\sqrt[3]{9\beta_1\beta_2}}{2}$$

Region B:

$$\lambda_1^* = 0 \quad , \quad \lambda_2^* = \frac{1}{\sqrt{6\beta_2}}$$

$$L(\lambda^*) = \sqrt{\frac{2}{3}\beta_2}$$

Region C:

$$\lambda_1^* = 0 \quad , \quad \lambda_2^* = \frac{3}{2} - 3\beta_2$$

$$L(\lambda^*) = \frac{3}{2}\beta_2^2 + \frac{3}{2}\beta_2 + \frac{1}{8}$$

Region D:

$$\lambda_1^* = \frac{1}{7}(36\beta_2 - 48\beta_1 + 6) \quad , \quad \lambda_2^* = \frac{1}{7}(36\beta_1 - 48\beta_2 + 6)$$

$$L(\lambda^*) = \frac{1}{7}(-24\beta_1^2 - 24\beta_2^2 + 36\beta_1\beta_2 + 6\beta_1 + 6\beta_2 + \frac{1}{2})$$

Region E : to obtain closed form equations for the values λ^*, $L(\lambda^*)$ for β lying in E, an equation of either fourth or eighth degree should be solved explicitly. Yet, numerical evaluation is possible through the direct solution of the problem $\min_{\lambda \in S} L(\lambda)$, using an appropriate nonlinear programming routine.

A picture of the surface $L(\lambda^*)$, defined over the (β_1, β_2) plane and evaluated either analytically or numerically, is presented in Figure 3.

By calculating $\beta_i - E\, \underline{a}_i \underline{x}^L(\lambda^*)$ $(i=1,2)$ for the regions A,B,C and D and using the remark at the end of Theorem 3.3, one can verify that in A and D both constraints are i.o. binding in expectation at the optimal integer solution, while in B and C this is true only for the tighter constraint. This corresponds to intuition: when β_2 is sufficiently small with respect to β_1, as in B and C, the first constraint can be disregarded, so that MK with two constraints is reduced to a simple knapsack problem. To support this conclusion, observe that the values of λ_2^* and $L(\lambda^*)$ obtained for the regions B and C are identical to the corresponding ones derived for the case m=1.

Similar calculations can be carried our for $m \geq 3$, even though in these cases only numerical approximation of λ^* and $L(\lambda^*)$ is possible for many values of β. The computation of $L(\lambda)$ and $\nabla L(\lambda)$ can be seen to amount to integrating the density function over regions defined by linear inequalities. In many situations, closed form expressions for these integrals can be derived in principle.

5. Probabilistic analysis of an approximation algorithm

The algorithm for solving the MK problem considered in this section belongs to the family of *generalized greedy* heuristics, that can be described as follows. Given m positive *weights* $y_1, y_2, \ldots, y_m$, order the items according to decreasing ratios

$$\rho_j = \frac{c_j}{\sum_{i=1}^{m} y_i a_{ij}} \qquad (j=1,2,\ldots,n).$$

The generalized greedy heuristic, denoted by G(y), then select items according to this order until the next item considered will violate one of the constraints if added to the knapsacks. Let $z_n^G(y)$ indicate the value of the solution obtained by this heuristic. Obviously, the quality of the approximation $z_n^G(y)$ is affected by the choice of the weight vector y. The behaviour of G(y) as a function of y will be analyzed in a forthcoming paper. In particular, it is possible to show that (under a non-degeneracy assumption) the choice $y=\lambda_n^*$, where for each problem instance λ_n^* is the vector of optimal dual multipliers, leads to a solution $z_n^G(\lambda_n^*)$ which is the same as the integer round-down of the optimal solution of MKLP. A drawback of this choice seems to be that y depends on the particular problem instance in a non-trivial way. Yet, in light of Theorem 3.4 the choice $y=\lambda^*$ where λ^* is defined (as in Section 3) to be the unique minumum of $L(\lambda)$, seems to be a reasonable alternative, at least in a probabilistic framework.

In fact, we shall show that the generalized greedy heuristic corresponding to weights $y=\lambda^*$ is asymptotically optimal wp1 with respect to the stochastic model of MK introduced in Section 2.

The following result providing bounds on the probability that a normalized sum of r.v.'s deviates from its mean is due to Hoeffding /8/ and will be useful in what follows.

<u>LEMMA 5.1</u>: Let $X_1, X_2, \ldots X_n$ be independent r.v.'s taking values in the interval $[0,1]$ and having finite mean and variance. Then, for $0<t<1-\frac{1}{n}E\left[\sum_{j=1}^{n} X_j\right]$,

$$\Pr\left\{\frac{1}{n}\left(\sum_{j=1}^{n} X_j - E\left[\sum_{j=1}^{n} X_j\right]\right) \geq t\right\} \leq e^{-2nt^2} . \tag{5.1}$$

$$\Pr\left\{\frac{1}{n}\left(E\left[\sum_{j=1}^{n} X_j\right] - \sum_{j=1}^{n} X_j\right) \geq t\right\} \leq e^{-2nt^2} , \tag{5.2}$$

$$\Pr\left\{\frac{1}{n}\left|\sum_{j=1}^{n} X_j - E\left[\sum_{j=1}^{n} X_j\right]\right| \geq t\right\} \leq 2e^{-2nt^2} . \blacksquare \tag{5.3}$$

Define the r.v.

$$\underline{\mu}_n = \inf\{\mu \geq 0 \mid \sum_{j=1}^{n} \underline{a}_{ij}\underline{x}_j^L(\mu\lambda^*) \leq b_i \quad (i=1,2,\ldots,m)\} .$$

Clearly, $\underline{\mu}_n$ can be interpreted as the value of the weight ρ_j corresponding to the last item included in the knapsacks by the algorithm $G(\lambda^*)$.

We first establish a convergence result for the sequence of r.v.'s $\{\underline{\mu}_n\}$.

THEOREM 5.1: $\underline{\mu}_n \xrightarrow{wp1} 1.$

Proof: The proof will be split in two parts. We first show that $\underline{\mu}_n \xrightarrow{wp1} \bar{\mu}$, where

$$\bar{\mu} = \inf\{\mu \geq 0: \ E\, \underline{a}_i \underline{x}^L(\mu\lambda^*) \leq \beta_i \quad (i=1,2,\ldots m)\}.$$

We then prove that $\bar{\mu} = 1$.

Let $\underline{I}_n(\mu)$ and $I(\mu)$ be respectively the indicator functions of the sets

$$\underline{\Omega}_n = \{\mu \geq 0: \sum_{j=1}^{n} \underline{a}_{ij} \underline{x}_j^L(\mu\lambda^*) \leq b_i \quad (i=1,2,\ldots,m)\}$$

and

$$\Omega = \{\mu \geq 0: E\, \underline{a}_i \underline{x}^L(\mu\lambda^*) \leq \beta_i \quad (i=1,2,\ldots m)\}.$$

We want to show that for $\mu \geq 0$, $\underline{I}_n(\mu) \xrightarrow{wp1} I(\mu)$. Assume first that there exists at least one index k, $1 \leq k \leq m$, for which $E\, \underline{a}_k \underline{x}^L(\mu\lambda^*) > \beta_k$, so that $I(\mu) = 0$. We then have that

$$\Pr\{|\underline{I}_n(\mu)-I(\mu)|>\varepsilon\}=\Pr\{\frac{1}{n}\sum_{j=1}^{n}\underline{a}_{ij}\underline{x}_j^L(\mu\lambda^*)\leq\beta_i\ (i=1,2,\ldots,m)\}$$

$$\leq \Pr\{\frac{1}{n}\sum_{j=1}^{n}\underline{a}_{kj}\underline{x}_j^L(\mu\lambda^*)\leq\beta_k\}$$

$$=\Pr\{\frac{1}{n}\sum_{j=1}^{n}\underline{a}_{kj}\underline{x}_j^L(\mu\lambda^*)-E\ \underline{a}_k\underline{x}^L(\mu\lambda^*)\leq\beta_k-E\ \underline{a}_k\underline{x}^L(\mu\lambda^*)\}$$

$$\leq e^{-2n(\beta_k-E\ \underline{a}_k\underline{x}^L(\mu\lambda^*))^2}, \qquad (5.4)$$

where the last inequality is derived from (5.2) by applying Lemma 5.1 to the i.i.d. r.v.'s $\underline{a}_{kj}\underline{x}_j^L(\mu\lambda^*)$ $(j=1,2,\ldots,n)$. Suppose now that $I(\mu)=1$. Then

$$\Pr\{|\underline{I}_n(\mu)-I(\mu)|>\varepsilon\}=\Pr\{\exists k:\ \frac{1}{n}\sum_{j=1}^{n}\underline{a}_{kj}\underline{x}_j^L(\mu\lambda^*)>\beta_k\}$$

$$\leq n\Pr\{\frac{1}{n}\sum_{j=1}^{n}\underline{a}_{kj}\underline{x}_j^L(\mu\lambda^*)-E\ \underline{a}_k\underline{x}^L(\mu\lambda^*)>\beta_k-E\ \underline{a}_k\underline{x}^L(\mu\lambda^*)$$

$$\leq ne^{-2n\ (\beta_k-E\ \underline{a}_k\underline{x}^L(\mu\lambda^*))^2}, \qquad (5.5)$$

again by Lemma 5.1.

Combining (5.4) and (5.5) it follows that the series $\Sigma\Pr\{|\underline{I}_n(\mu)-I(\mu)|>\varepsilon\}$ is convergent for $\mu\geq 0$ and any $\varepsilon>0$, so that $\underline{I}_n(\mu)\xrightarrow{wp1} I(\mu)$.

Turning now to the sequence $\underline{\mu}_n$, we have that

$$\Pr\{|\underline{\mu}_n-\bar{\mu}|>\varepsilon\} \leq \Pr\{\underline{\mu}_n-\mu>\varepsilon\} + \Pr\{\bar{\mu}-\underline{\mu}_n>\varepsilon\}$$

$$\leq \Pr\{I(\bar{\mu}+ \frac{\varepsilon}{2})-\underline{I}_n(\bar{\mu}+ \frac{\varepsilon}{2})=1\} + \Pr\{\underline{I}_n(\bar{\mu}- \frac{\varepsilon}{2})-I(\bar{\mu}- \frac{\varepsilon}{2}) = 1\}$$

for any $\varepsilon>0$, so that $\underline{\mu}_n \xrightarrow{wp1} \bar{\mu}$.

Now, to show that $\bar{\mu}=1$, we recall condition (ii) of Lemma 3.3,

$$E\ \underline{a}_i\underline{x}^L(\lambda^*) \leq \beta_i \qquad (i=1,2,\ldots,m), \tag{5.6}$$

noticing that it implies that $\bar{\mu}\leq 1$. Suppose now that $\bar{\mu}<1$. Since the functions $E\ \underline{a}_i\underline{x}^L(\lambda^*)$ $(i=1,2,\ldots,m)$ are strictly decreasing in each component λ_k $(k=1,2,\ldots,m)$ (see the Appendix), $\bar{\mu}<1$ implies that all inequalities in (5.6) hold strictly. This in turn implies by condition (i) of Lemma 3.3 that $\lambda^*=0$, contradicting Lemma 3.2. ∎

We are now in a position to prove the final result.

<u>THEOREM 5.2</u>: $\underline{z}_n^G(\lambda^*) \xrightarrow{wp1} L(\lambda^*)$.

<u>Proof</u>: By definition of $\underline{\mu}_n$, it follows that

$$\underline{z}_n^G(\lambda^*) = \sum_{j=1}^{n} \underline{c}_j\underline{x}_j^L(\underline{\mu}_n\lambda^*).$$

Let $\varphi(\underline{\mu}_n)=E\left[\underline{c}_1\underline{x}_1^L(\underline{\mu}_n\lambda^*)\ |\ \underline{\mu}_n\right]$. We have that, for $\varepsilon>0$,

$$\Pr\{|\frac{1}{n}\underline{z}_n^G(\lambda^*)-L(\lambda^*)|>\varepsilon\} = \Pr\{|\frac{1}{n}\sum_{j=1}^{n}\underline{c}_j\underline{x}_j^L(\underline{\mu}_n\lambda^*)-L(\lambda^*)|>\varepsilon\}$$

$$\leq \Pr\{|\frac{1}{n}\sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\underline{\mu}_n\lambda^*) - \underline{\varphi}(\underline{\mu}_n)| + |\underline{\varphi}(\underline{\mu}_n) - L(\lambda^*)| > \varepsilon\}$$

$$\leq \Pr\{|\frac{1}{n}\sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\underline{\mu}_n\lambda^*) - \underline{\varphi}(\underline{\mu}_n)| > \frac{\varepsilon}{2}\} + \Pr\{|\underline{\varphi}(\underline{\mu}_n) - L(\lambda^*)| > \frac{\varepsilon}{2}\} . \quad (5.7)$$

The function $\varphi(\mu)$ is continuous, so that $\underline{\varphi}(\underline{\mu}_n) \xrightarrow{wp1} L(\lambda^*)$ by Theorem 5.1. It follows that

$$\Sigma \Pr\{|\underline{\varphi}(\underline{\mu}_n) - L(\lambda^*)| > \frac{\varepsilon}{2}\} < \infty. \quad (5.8)$$

Moreover, letting $F_n(\mu)$ be the distribution of $\underline{\mu}_n$, we have that

$$\Pr\{|\frac{1}{n}\sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\underline{\mu}_n\lambda^*) - \underline{\varphi}(\underline{\mu}_n)| > \frac{\varepsilon}{2}\} =$$

$$= \int \Pr\{|\frac{1}{n}\sum_{j=1}^{n} \underline{c}_j \underline{x}_j^L(\mu\lambda^*) - \underline{\varphi}(\mu)| > \frac{\varepsilon}{2}\} \, dF_n(\mu)$$

$$\leq 2e^{-2n\frac{\varepsilon^2}{4}} , \quad (5.9)$$

the last inequality being derived from (5.3), since Lemma 5.1 can be applied to the i.i.d. r.v.'s $\underline{c}_j \underline{x}_j^L(\mu\lambda^*)$.

Combining (5.7), (5.8) and (5.9), one finally obtains the required result. ■

Further investigation of the family of heuristics G(y) seems to be an interesting topic for future research. Although the results of this paper carry over immediately to the minimization of MK (with $\geq$ replacing $\leq$), preliminary investigations suggest that the heuristic's worst case behaviour for these two models differs substantially.

/1/ G. d'Atri, Probabilistic analysis of the knapsack problem. Technical Report N. 7, Groupe de Recherche 22, Centre National de la Recherche Scientifique, Paris, 1978.

/2/ G. Ausiello, A. Marchetti-Spaccamela and M. Protasi, Probabilistic analysis of the solution of the knapsack problem. Proc. 10th IFIP Conference, Springer-Verlag, New York, 1982, 557-565.

/3/ G. Dobson, Worst-case analysis of greedy heuristics for integer programming with nonnegative data. Math. Oper. Res. 4 (1982), 515-531.

/4/ A.V. Fiacco and G.P. McCormick, Non linear programming: sequential unconstrained minimization techniques. Wiley, New York, 1968.

/5/ A.M. Frieze and M.R.B. Clarke, Approximation algorithms for the m-dimensional 0-1 knapsack problem:worst-case and probabilistic analysis. Eur. J. Oper. Res. 15 (1984), 100-109.

/6/ M.R. Garey and D.S. Johnson, Computers and intractability: guide to the theory of NP-completeness. Freeman, San Francisco, 1978.

/7/ A.V.Goldberg and A. Marchetti-Spaccamela, On finding the exact solution of a 0-1 knapsack problem. Proc. 16th Ann. ACM Symp. on Theory of Computing, Association for Computing Machinery, New York, 1984, 359-368.

/8/ W. Hoeffding, Probability inequalities for sums of bounded random variables. Am. Stat. Ass. J. 3 (1963), 13-30.

/9/ B. Korte and R. Schrader, On the existence of fast approximation schemes. Report No. 80163-OR, Institut fur Okonometrie und Operations Research, Rheinische Friederich Wilhelms Universitat, Bonn, 1980.

/10/ J. Lorie and L. Savage, Three problems in capital rationing. Journal of Business 28 (1955), 229-239.

/11/ G.S. Lueker, On the average difference between the solution to linear and integer knapsack problems. Applied Probability-Computer Science, the Interface, Vol. 1, Birkhauser, 1982.

/12/ R.I. Rockafellar, Convex Analysis. Princeton University Press, Princeton, 1970.

/13/ L. Schwartz, Cours d'analyse, Vol. 1. Hermann, Paris, 1967.

Appendix

The purpose of this section is that of proving Lemmas 3.2 and 3.3. To this end, the following two results are useful.

LEMMA A.1: The functions $E\,\underline{cx}^L(\lambda)$, $E\,\underline{a_j x}^L(\lambda)$, $j=1,2,\ldots m$, are continuously differentiable and strictly decreasing with respect to each component λ_k $(k = 1,2,\ldots m)$.

Proof: It will be shown that the partial derivatives $\dfrac{\partial E\underline{a_j x}^L(\lambda)}{\partial \lambda_k}$ exist and are continuous and strictly negative, for each $\lambda_k \geq 0$. This implies the required result for $E\underline{a_j x}^L(\lambda)$; similar arguments can be applied to prove the Lemma also for $E\underline{cx}^L(\lambda)$.

To simplify the notations, assume that k=m without loss of generality. Let $a'=(a_1,a_2,\ldots a_{m-1})$, and define

$$D = \{a,c \mid 0\leq a\leq 1, \quad 0\leq c\leq 1, \quad c\geq \sum_{i=1}^{m} \lambda_i a_i\}$$

$$D' = \{a',c \mid 0\leq a'\leq 1, \quad 0\leq c\leq 1, \quad c\geq \sum_{i=1}^{m-1} \lambda_i a_i\}$$

$$D_m = \{a_m \mid 0\leq a_m\leq 1, \quad \lambda_m a_m \leq c - \sum_{i=1}^{m-1} \lambda_i a_i\} \quad .$$

Let $F(a_1,\ldots a_m,c) = F_c(c)\cdot \prod_{i=1}^{m} F_i(a_i)$ and $F'(a_1,\ldots a_m,c) = F_c(c)\cdot \prod_{i=1}^{m-1} F_i(a_i)$. By Fubini's theorem, we have that

$$E\underline{a}_i\underline{x}^L(\lambda) = \int_D a_j dF = \int_{D'} \left[\int_{D_m} a_j dF_m\right] dF' = \int_{D'} U(\lambda_m) dF' \qquad (A.1)$$

where

$$U(\lambda_m) = \int_{D_m} a_j dF_m = \int_0^{\gamma(\lambda_m)} a_j dF_m,$$

with

$$\gamma(\lambda_m) = \begin{cases} (c - \sum_{i=1}^{m-1} \lambda_i a_i)/\lambda_m & \text{if } c - \sum_{i=1}^{m-1} \lambda_i a_i \leq \lambda_m \\ 1 & \text{otherwise} . \end{cases}$$

For $0<\lambda_m<c-\sum_{i=1}^{m-1}\lambda_i a_i$, the function $U(\lambda_m)$ is constant, hence it is continuously differentiable.

For $\lambda_m > c-\sum_{i=1}^{m-1}\lambda_i a_i$, by the absolute continuity of F_m, it follows that $U(\lambda_m)$ is differentiable, and

$$\frac{dU}{d\lambda_m} = \begin{cases} -a_j f_m\left(\dfrac{c-\sum_{i=1}^{m-1}\lambda_i a_i}{\lambda_m}\right)\dfrac{c-\sum_{i-1}^{m-1}\lambda_i a_i}{\lambda_m^2}, & j \neq m \\[2ex] -f_m\left(\dfrac{c-\sum_{i=1}^{m-1}\lambda_i a_i}{\lambda_m}\right)\dfrac{\left(c-\sum_{i=1}^{m-1}\lambda_i a_i\right)^2}{\lambda_m^3}, & j = m . \end{cases} \qquad (A.2)$$

For each $\lambda_m>0$, the function $\frac{dU}{d\lambda_m}$ is continous for all (a',c) in D' except on the hyperplane

$c-\sum_{i=1}^{m-1}\lambda_i a_i = \lambda_m$. Therefore, the function $U(\lambda_m)$ is continuously differentiable for $\lambda_m>0$ almost everywhere in D' with respect to dF'. Since D' does not depend on λ_m, it can be concluded /13/ that

$\dfrac{\partial E\,\underline{a}_j\underline{x}^L(\lambda)}{\partial\lambda_m}$ exists and is continuous for $\lambda_m>0$. In addition,

$$\frac{\partial E\underline{a}_j\underline{x}^L(\lambda)}{\partial\lambda_m} = \int_{D'} \frac{dU}{d\lambda_m}\,dF' = \int_{D'_o} \frac{dU}{d\lambda_m}\,dF', \qquad (A.3)$$

where $D'_o=\{a',c\colon 0\le a'\le 1,\ 0\le c\le 1,\ 0\le c-\sum_{i=1}^{m-1}\lambda_i a_i \le \lambda_m\}$,

since $\dfrac{dU}{d\lambda_m} = 0$ on $D'-D'_o$.

It remains to be shown that the right derivative of $E\underline{a}_j\underline{x}^L(\lambda)$ in $\lambda_m=0$ exists and is equal to the limit of (A.3) for $\lambda_m\to 0^+$. To see this, define

$$D'_a = \{a' \mid 0\le a'\le 1,\ \sum_{i=1}^{m-1}\lambda_i a_i \le 1\},$$

$$D'_c = \{c \mid 0\le c\le 1,\ 0\le c-\sum_{i=1}^{m-1}\lambda_i a_i \le \lambda_m\},$$

$$\bar{D}_a = \{a \mid 0\le a\le 1,\ \sum_{i=1}^{m-1}\lambda_i a_i \le 1\}.$$

We have, for $\lambda_m>0$,

$$\lim_{\lambda_m\to 0^+} \frac{\partial E\underline{a}_j\underline{x}^L(\lambda)}{\partial\lambda_m} = \lim_{\lambda_m\to 0^+} \int_{D'_a} \left[\int_{D'_c} \frac{dU}{d\lambda_m}\,dF_c\right] dF'_a, \qquad (A.4)$$

where $F'_a(a') = \prod_{i=1}^{m-1} F_i(a'_i)$. Applying the substitution $c = \sum_{i=1}^{m} \lambda_i a_i$ to the inner integral in (A.4), one obtains from (A.2), and for all j=1,2,...m,

$$\int_{D'_c} \frac{dU}{d\lambda_m} dF_c = \int_0^1 - a_j a_m \; f_c\left(\sum_{i=1}^{m} \lambda_i a_i\right) dF_m \; .$$

Since $\bar{D}_a$ does not depend on λ_m, we have

$$\lim_{\lambda_m \to 0^+} \frac{\partial E\underline{a}_j\underline{x}^L(\lambda)}{\partial \lambda_m} = \int_{\bar{D}_a} \lim_{\lambda_m \to 0^+} \left[-a_j a_m f_c\left(\sum_{i=1}^{m} \lambda_i a_i\right)\right] dF_a$$

$$= \int_{\bar{D}_a} -a_j a_m f_c\left(\sum_{i=1}^{m-1} \lambda_i a_i\right) dF_a \; , \qquad \text{(A.5)}$$

where $F_a(a) = \prod_{i=1}^{m} F_i(a_i)$.

It can be easily verified that the right derivative of $E\underline{a}_j\underline{x}^L(\lambda)$ in $\lambda_m = 0$ exists and is equal to the last term in (A.5).

To conclude that $E\underline{a}_j\underline{x}^L(\lambda)$ is strictly decreasing with respect to λ_m, observe that

$\frac{\partial E\underline{a}_j\underline{x}^L(\lambda)}{\partial \lambda_m}$ is strictly negative for $\lambda_m \geq 0$. Indeed, it follows from (A.2), (A.3) and (A.5) that this derivative is equal, for every $\lambda_m \geq 0$, to the integral of a strictly negative function over a set of positive measure. ■

LEMMA A.2: For k=1,2,...m, the following relation holds:

$$\frac{\partial E\ \underline{c}\underline{x}^L(\lambda)}{\partial \lambda_k} = \sum_{i=1}^{m} \lambda_i \frac{\partial E\ \underline{a}_i \underline{x}^L(\lambda)}{\partial \lambda_k} .$$

Proof: Again, without loss of generality, assume k=m.

Define

$$D_a = \{a:\ 0 \le a \le 1,\quad \sum_{i=1}^{m} \lambda_i a_i \le 1\} ,$$

$$D_c = \{c:\ 0 \le c \le 1,\quad c \ge \sum_{i=1}^{m} \lambda_i a_i\} .$$

We have, by Lemma A.1,

$$\frac{\partial E\ \underline{c}\underline{x}^L(\lambda)}{\partial \lambda_m} = \frac{\partial}{\partial \lambda_m} \left(\int_{D_a} \left[\int_{D_c} c dF_c \right] dF_a \right) =$$

$$= \int_{D_a} \frac{dU_c}{d\lambda_m} dF_a , \tag{A.6}$$

where $U_c(\lambda_m) = \int_{\sum_{i=1}^{m} \lambda_i a_i}^{1} c dF_c = E\left[c\right] - \int_{0}^{\sum_{i=1}^{m} \lambda_i a_i} c dF_c .$

Hence, for $\lambda_m > 0$, $\frac{dU_c}{d\lambda_m} = - \sum_{i=1}^{m} \lambda_i a_i a_m\ f_c(\sum_{i=1}^{m} \lambda_i a_i)$, which,

substituted in (A.6), gives

$$\frac{\partial E\ \underline{c}\underline{x}^L(\lambda)}{\partial \lambda_m} = \sum_{i=1}^{m} \lambda_i \int_{D_a} -a_i a_m f_c\ (\sum_{i=1}^{m} \lambda_i a_i) dF_a . \tag{A.7}$$

Applying the substitution $a_m=(c-\sum_{i=1}^{m-1}\lambda_i a_i)/\lambda_m$ to the integral in (A.7) and performing straightforward calculations, we obtain

$$\int_{D_a} -a_i a_m f_c \left(\sum_{i=1}^{m}\lambda_i a_i\right) dF_a = \int_{D'} \frac{dU}{d\lambda_m}\, dF' , \qquad (A.8)$$

where D', F' and U are defined as in Lemma A.1. Combining (A.7) and (A.8), the required result easily follows for $\lambda_m>0$. Continuity of the partial derivatives implies that Lemma A.2 holds also for $\lambda_m=0$. ■

Lemma 3.2 and 3.3 can now be proved.

Proof of Lemma 3.2: For k=1,2,...m, we have

$$\frac{\partial L}{\partial \lambda_k} = \beta_k - \frac{\partial E\, \underline{a}_k\underline{x}^L(\lambda)}{\partial \lambda_k} + \frac{\partial E\, \underline{cx}^L(\lambda)}{\partial \lambda_k} - \sum_{i=1}^{m}\lambda_i \frac{\partial E\, \underline{a}_i\underline{x}^L(\lambda)}{\partial \lambda_k}$$

$$= \beta_k - \frac{\partial E\, \underline{a}_k\underline{x}^L(\lambda)}{\partial \lambda_k} ,$$

by Lemmas A.1 and A.2. This implies that $L(\lambda)$ is continuously differentiable and that $\nabla L(\lambda)=\beta-E\,\underline{ax}^L(\lambda)$.

Moreover, we have

$$\frac{\partial}{\partial \lambda_j}\left(\frac{\partial L}{\partial \lambda_k}\right) = \frac{\partial}{\partial \lambda_j}\left(\beta_k - E\, \underline{a}_k\underline{x}^L(\lambda)\right) =$$

$$= - \frac{\partial E\, \underline{a}_k\underline{x}^L(\lambda)}{\partial \lambda_j} > 0 ,$$

again by Lemma A.1. This implies that $L(\lambda)$ is twice continuously differentiable and strictly convex over the region $\lambda\geq 0$, as its Hessian matrix $H(\lambda)$ is positive definite for $\lambda \geq 0$.

$L(\lambda)$ is strictly convex, so that if it has minimum over the region $\lambda \geq 0$, that minimum must be unique. To see that at least one minimum exists, one can proceed as in the proof of Lemma 3.4.

<u>Proof of Lemma 3.3:</u> Since $L(\lambda)$ is continuously differentiable, and the constraints $\lambda > 0$ do satisfy the first-order constraint qualifications in λ^* /4/, Kuhn-Tucker conditions for optimality hold at λ^* and lead immediately to (i) and (ii).

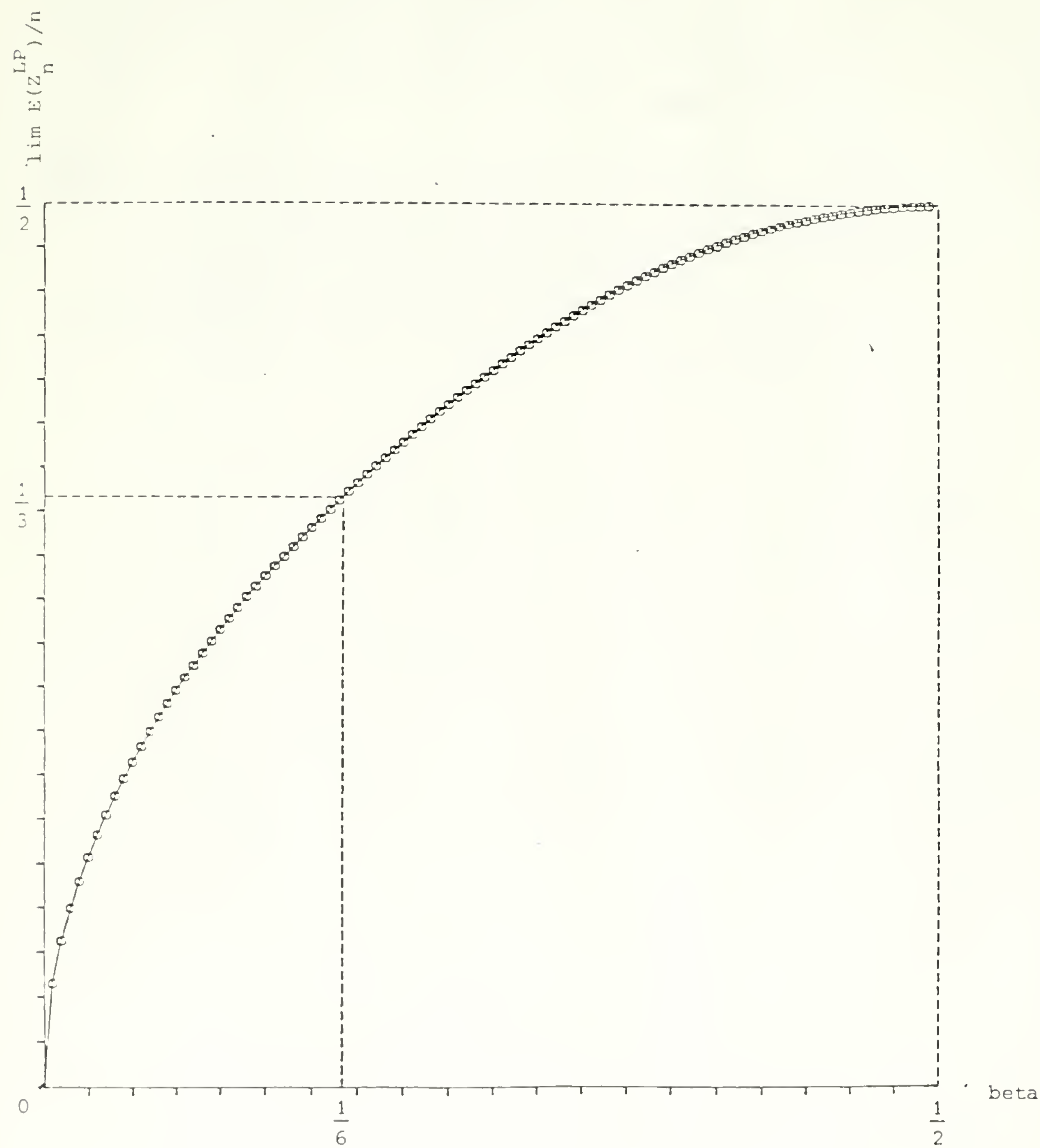

Figure 1

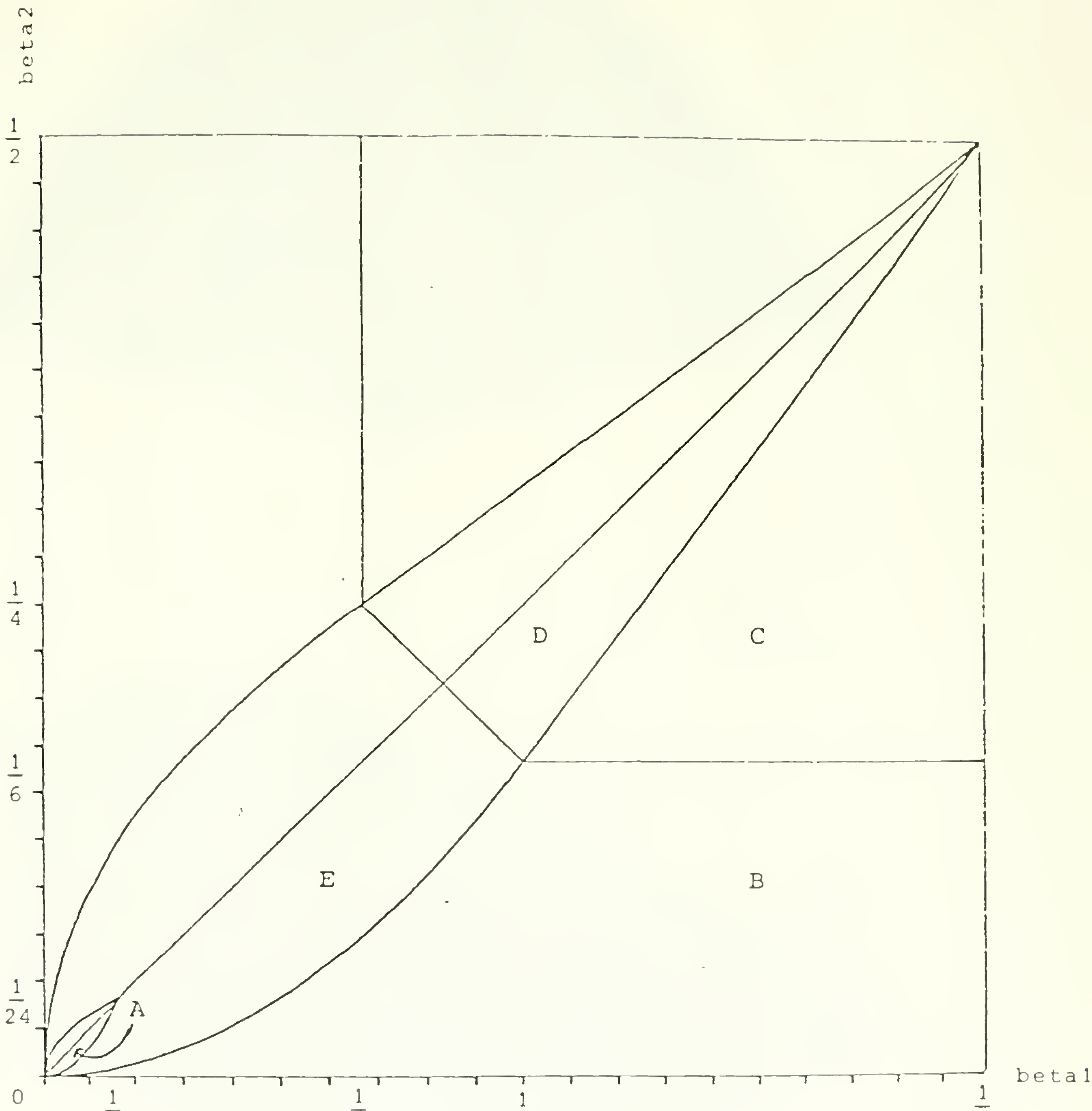

Figure 2

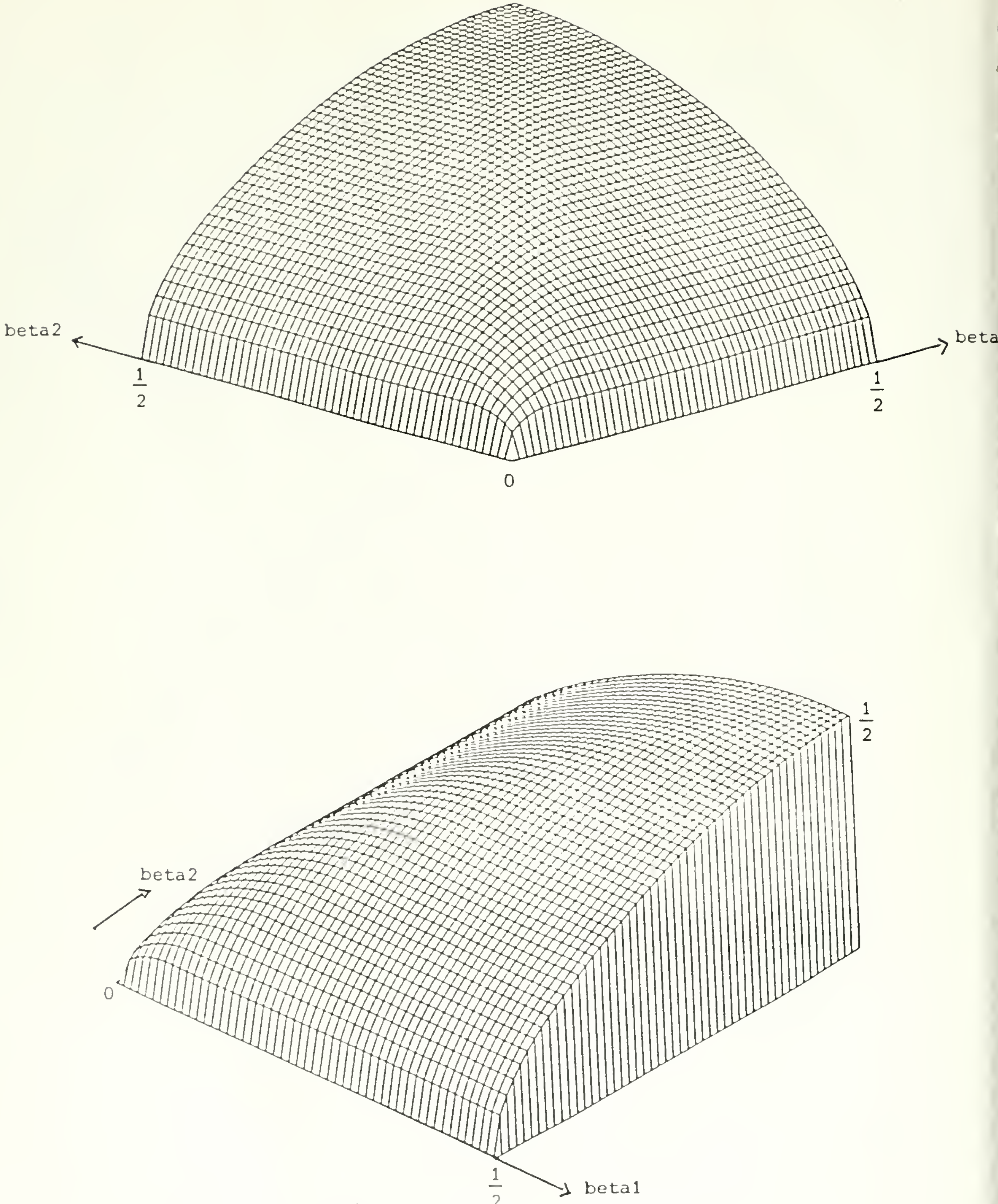

Figure 3

Bar Code
on Last
Page